WO 97/15008 P2

(2)

[19]中华人民共和国专利局

[51]Int.Cl[6]
G06F 11/00

[12] 发明专利申请公开说明书

[21] 申请号 96190606.5

[43]公开日 1997 年 9 月 10 日

[11] 公开号 CN 1159234A

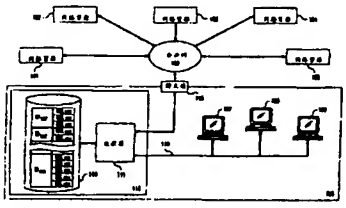
[22]申请日 96.6.6
[30]优先权
[32]95.6.6 [33]US[31]08/469,342
[86]国际申请 PCT/US96/09510 96.6.6
[87]国际公布 WO97/15008 英 97.4.24
[85]进入国家阶段日期 97.2.4
[71]申请人 美国电报电话IPM公司
地址 美国佛罗里达
[72]发明人 布林达·苏·贝克
艾里斯·格劳瑟
[74]专利代理机构 中国国际贸易促进委员会专利商标事务所
代理人 范本国

权利要求书 2 页 说明书 7 页 附图页数 2 页

[54]发明名称 数据库访问控制的系统和方法

[57]摘要

用于有选择地控制数据库访问的系统和方法，这种系统和方法允许网络管理员或经营者限制特定系统用户访问某些公共或者说未受控制的数据库（即，WWW 和 Internet）上的信息。本发明采用了一个相关数据库来确定访问权限，并且管理员可以很容易地更新和修改该数据库。在该相关数据库中，对具体的资源标识号（即 URL）进行分类使之属于特定的访问组。相关数据库的安排应使得对系统中的每个用户而言，对特定资源的请求只有在该资源标识号在一个经管理员的特许用户能访问的访问组的情况下，该请求才从本地网络送到一个服务器上，由该服务器提供到该公共/未受控制的数据库的链路。在一个优选实施例中，本发明作为用户局域网中代理服务器的一部分而实现。



权 利 要 求 书

1.一种有选择地控制网络访问一个或多个资源的方法，包括:

一个相关数据库，包括一个存储的用户标识码及资源标识号列表，其中每个所述资源标识号对应一个或多个可通过网络访问的资源，所述存储的列表使每个所述用户标识码与一个或多个所述资源标识号相关;

一个处理器，用于接收要求通过网络访问一个或多个特定网络资源的请求，所述请求包括用户标识码，所述处理器还用于查询所述相关数据库，并根据所述存储的列表执行要求网络访问所述一个或多个特定网络资源的所述请求，所述存储的列表表明所接收的用户标识码与至少一个对应于所述一个或多个特定网络资源的资源标识号是否相关.

2.如权利要求 1 的发明，其中所述处理器被编程以在所述存储的列表表明所述接收到的用户标识码与至少一个对应于所述一个或多个特定网络资源的资源标识号相关的情况下执行所述访问请求.

3.如权利要求 1 的发明，其中所述处理器被编程以在所述存储的列表表明所述接收到的用户标识码与至少一个对应于所述一个或多个特定网络资源的资源识别号相关的情况下拒绝执行所述访问请求.

4.如权利要求 1 的发明，其中所述处理器包括在一个网络代理服务器中.

5.如权利要求 1 的发明，其中到所述一个或多个网络资源的访问通过公共网来进行.

6.如权利要求 1 的发明，其中每个所述用户标识码标识一个或多个用于通过网络访问一个或多个特定网络资源的终端.

7.如权利要求 1 的发明，其中每个所述用户标识码标识一个或多个被授权访问一个或多个特定网络资源的个体.

8.如权利要求 1 的发明，其中每个所述资源标识号对应于一个或多个用于访问一个或多个特定网络资源的统一资源定位器.

9.一种用于有选择地控制网络访问一个或多个特定资源的方法，包括步骤:

接收一个要求访问一个或多个特定网络资源的请求，其中所述请求

包括一个用户标识码和至少一个资源标识号；

将所述接收到的访问请求与一个包含一个用户标识码和资源标识号的存储列表的相关数据库比较，其中每个所述资源标识号对应于一个或多个可通过网络访问的资源，并且所述存储的列表使每个所述用户标识码与一个或多个所述资源标识号相关；

根据所述存储的列表来执行所述要求通过网络访问一个或多个特定网络资源的请求，所述存储的列表表明所述接收到的用户标识码和至少一个对应于所述一个或多个特定网络资源的资源标识号是否相关。

10.如权利要求9的方法，其中在假如所述存储的列表表明所述接收到的用户标识码与至少一个对应于所述一个或多个特定网络资源的资源标识号相关的情况下完成所述访问请求的执行。

11.如权利要求9的方法，其中在假如所述存储的列表表明所述接收到的用户标识码与至少一个对应于所述一个或多个特定网络资源的资源标识号相关的情况下拒绝所述访问请求的执行。

12.如权利要求9的方法，其中到所述一个或多个特定资源的所述网络访问是通过公共网来进行。

13.如权利要求9的方法，其中每个所述用户标识码标识一个或多个用于通过网络访问一个或多个特定网络资源的终端。

14.如权利要求9的方法，其中每个所述用户标识码标识一个或多个被授权能访问一个或多个特定网络资源的个体。

15.如权利要求9的方法，其中每个所述资源标识号对应于一个或多个用于访问所述一个或多个特定网络资源的统一资源定位器。

说　　明　　书

数据库访问控制的系统和方法

本发明涉及数据库访问的控制，尤其涉及根据另外的公共数据库有选择地提供这种控制.

通过著名的 Internet 网络的汇集，全球范围内的计算机上的文件和其它资源都可以被别的计算机用户公开地享用. 所有这些通过用超级文本标记语言（ Hypertext Mark-up Language ）（ “ HTML ” ）写成的文件链接起来的可公用资源的汇集就是著名的万维网（ World Wide Web ）（ “ WWW ” ）.

连在 Internet 上的计算机用户可使一个称作委托程序的程序请求 WWW 的部分资源. 然后服务器程序处理请求并返回所指定的资源（假设正好有）已经采用了一种叫作统一资源定位器（ “ URL ” ）（ Uniform Resource Locator ）的标准命名约定，该约定包括几种类型的位置名称，当前包括诸如超级文本传输协议（ “ http ” ），文件传输协议（ “ ftp ” ），地鼠（ gopher ）和广域信息服务（ “ WAIS ” ）等子类. 当资源被卸载下来时，它可以包括附加资源的 URL. 这样委托程序的用户就很容易知道他或她未专门请求的新资源的存在.

通过 WWW 可访问的各种各样的资源是在全世界各地的计算机上由许多不同的人来创建和维护的，还没有对这些内容进行集中控制. 由于包含在这些未受控制的信息汇集中的特定类型的信息或图像对某些用户可能是不合适的，因此期望有选择地限制对 WWW 资源的访问. 例如，父母们或学校老师们可能希望让小孩去访问有用的信息，而不是去访问淫秽的东西（小孩可能由于对 WWW 天真探索或偶然的 URL 的卸载而接触到这些东西）. 另一种情况是学校老师想在课堂上只让他们的学生访问某一组特定资源. 第三种情况是商业人士可能只想让他们的雇员仅访问与工作有关的资源，而不想让他们在其它 WWW 探索上浪费时间. 一般来讲，可能需要限定一个特定的用户在不同的时间里访问不同的资

源，如同学生在不同课程的课堂上应限于不同组的资源一样。

有些部门（例如学校）要求使用者遵守限制 WWW 探索的条例，例如同意不卸载下淫秽的材料。然而，对这种条例的自愿遵守并不能防止意外地卸下本不该卸下的资源，因为在卸载下并经浏览前不易识别出这些资源是受禁止的。

自然，出现了诸如“防火墙”等技术方案以限制或阻碍对 WWW 和 Internet 的访问。这些防火墙是基于软件的网关，它们被共同安装以保护局域网（“LAN”）的计算机不受外来者的攻击。安装防火墙的一个后果是 WWW 的客户不再能直接与 WWW 服务器接触。通常，这种限制太严格了，使用户求助于由 WWW 的客户机能直接接触的“代理服务器”。这种代理服务器有能力透过防火墙而提出请求，从而提供与 Internet 服务器的通信。为了提高效率，代理服务器还可能在本地高速缓存一些资源。当前的客户机和代理服务器能对 WWW 上的所有公共资源进行访问，因为 WWW 资源的设计本身就不是让特定的用户只能请求某些资源，而阻止该用户对另外资源的访问。

在提供间接访问的系统中，对可用 WWW 资源的“过滤”可能会有作用。在这些系统中，信息提供者将从 WWW 上卸载下资源，并维持这些资源的拷贝。用户则对这些拷贝进行访问。信息提供者能在从 WWW 上获得资源时浏览这些资源，并在把这些资源提供给用户前在编辑过程中删除任何不适宜的或淫秽的东西。这种方案的缺点是，与 WWW 上的原始资源相比，由信息提供者所提供的材料可能已过时了。

在对 WWW 资源进行“过滤”访问的另一种方案中，代理服务器给用户提供允许访问的资源的菜单，用户可获取通过一系列菜单资源链接能达到的任何资源。仅允许用户通过该菜单请求 URL。这种方法有两个缺点。第一点是很多材料即使本身是可被接受的，但因为它们包含与不适宜的材料的链接而不得不从菜单中取下来。第二点是某种资源可能因随时间变化而包括了可能导向到不适宜的材料的新链路，从而不经意地给用户提供了一条对这些材料的访问。

在对 WWW 资源进行“过滤”访问的另一方案中，客户机或代理服务器检查每一资源上是否有不允许的词（如，猥亵淫秽的或性术语方面

的等），并只把那些不包含这类词的资源显示给用户．然而，这种方法不能过滤掉图象，也不能禁止掉那些不包含这些特定词但内容仍然是不适宜的资源．

还有一种防止用户接触到不适宜的或淫秽材料的方法已由计算机或视频游戏的制造者提供．他们自愿标定游戏中暴力、裸体/性和语言程度的等级．虽然在 WWW 中还未采用这种协定，但同样地也可以给 WWW 资源定等级，可能的话并加入数字标记以防伪造．从而可以对 WWW 客户机进行编程，使之对那些未定级或给定等级是观众不能接受的资源不进行存储或显示．这种方法的缺点是需要说服那么多提供应用服务器（通常基于非职业或业余（ pro bono ））的人与定级小组合作．

现有这些限制用户对未受控制的公用数据库资源（例如在 WWW 上的数据库资源）的访问的系统都有明显的缺点．目前为止，还没有一种简单的方法能让管理人员（即老师、监护人、系统管理人员等）有选择地控制一个或多个使用者对 WWW 的访问，而又不显著妨碍用户与 Internet 交流的能力．

本发明克服了现有的有选择地控制数据库访问的方案的缺点，这是通过提供一种允许网络管理员或经营者限制特定系统的用户对某些公用的或者说未受控制的数据库（即 WWW 和 Internet ）信息的访问的系统和方法来实现的．本发明采用了一个相关数据库来确定访问权限，并且该数据库可以很容易地被管理员更新和修改．在相关数据库中，把具体的资源标识号（即 URL ）进行分类使之属于特定访问组．相关数据的安排应使得对系统的每一用户而言，用户对特定资源的请求只有在该资源标识号在一个经管理员的特许用户已可以访问的组中时才通过局域网传到一个服务器上，由该服务器提供一条到公共/未受控制数据库的链路．在一个优选实施例中，本发明作为用户局域网中代理服务器的一部分而实现．

附图中，

图 1 是实现本发明的典型系统的简图；及

图 2 是另一幅简图，说明了为便于用户/用户终端类别识别而采用的图 1 所示系统的另一种安排．

图 1 是一个实现本发明的典型系统的简图. 如图所示, 系统包括公共网 100、网络资源 101 - 105 和用户站 106. 用户站 106 处的特定用户可通过用户终端 107、108 和 109 访问公共网 100. 每个用户终端通过局域网(“LAN”) 110 与代理服务器 112 内的处理器 111 相连. 最后, 代理服务器 112 通过防火墙 113 提供处理器 111 到公共网 100 的连接.

要求通过公共网的 100 访问网络资源(101 - 105)的用户终端 107 - 109 的请求被传送到代理服务器 112 内的处理器 111. 在本发明的这一具体实施例中, 假设所提交的请求具有 URL 形式. 如同在现有技术中是公知的那样, 当 URL 提交到代理服务器上时, 通过把标识头装到该 URL 上, 代理服务器就可识别出发该请求的特定用户终端. 对图 1 所示系统而言, 用户终端 107 的标识码是 ID_{107}, 用户终端 108 的标识码是 ID_{108}, 以及用户终端 109 的标识码是 ID_{109}. 此外在图 1 系统中命名为 URL_{101}、URL_{102}、URL_{103}、URL_{104} 和 URL_{105} 的 URL 分别表示请求从网络资源 101、102、103、104 和 105 获取信息的请求.

一接收到进入的 URL, 处理器 111 程序就编程以确定来自 URL 头的请求用户终端的标识. 处理器 111 接着又利用该标识信息对接收到的 URL 和存在相关数据库 114 中的信息进行相互参照. 相关数据库 114 包括一个用户终端标识码(ID_{107}, ID_{108} …… ID_{109})的列表, 每个都与一个或多个 URL 标记相连. 该相关列表说明了可以从给定的用户终端发送到访问网络资源的特殊的 URL. 如图所示, 允许用户终端 107 使用的 URL 是 URL_{101}, URL_{102} 和 URL_{103}; 允许用户终端 108 使用的 URL 是 URL_{102} 和 URL_{104}; 允许用户终端 109 使用的 URL 是 URL_{101}, URL_{102}, URL_{103}, URL_{104} 和 URL_{105}. 存储在相关数据库 114 中的信息可以由用户站 106 处的本地管理者控制(即由系统管理员或有权决定特定用户能访问哪些 URL 的站点监护人控制).

在图 1 的系统中, 当请求用户终端发送一个与相关数据库 114 中特定终端识别码相关的 URL 时, 对由该 URL 所表示的信息的请求被送到公共网 100. 例如, 一从用户终端 107 接收到请求从网络资源 102 获取信息的请求, 处理器 111 就会访问相关数据库 114, 从而确定出 URL_{102}

确实是一个可允许的请求．确定出来后，处理器 111 将通过防火墙 113 把 URL_{102} 送到公共网 100．相反，如果处理器 111 收到的是一个与相关数据库 114 内请求终端标识码无关的 URL 时，对信息的请求被拒绝．例如，假如处理器 111 从用户终端 107 收到 URL_{104}，则访问相关数据库 114．由于 URL_{104} 不属于相关数据库 114 内与用户终端标识码 ID_{107} 相关的任一 URL，处理器 111 拒绝对信息的请求，且不向公共网 100 发送 URL．

在上面描述的特定实施例中，相关数据库 114 存储一个用户终端标识码及允许每个终端发送到公共网 100 上的各种 URL 的列表．应该理解的是可以对本发明进行修改，从而将与给定用户终端标识码相关的该相关 URL 用作一个该特定用户终端不允许连接的 URL 的列表．这种限定性列表功能可以通过再编程处理器 111 而很容易地实现．另外，还可以对本发明进行修改，从而由处理器 111 识别并存储在相关数据库 114 内的标识码是针对用户的，而不是针对用户终端的．换句话说，可以把系统修改成让系统通过一个私有口令或另外的标识码来识别出使用终端的特定个人．从而系统根据个人标识号来允许或拒绝特定 URL 的传输，而不管他们所使用的特定终端是什么．

还可对本发明代理服务器内的处理器和相关数据库进行修改以识别用户和/或用户终端类．在某一特定用户站中，所访问用户服务器的可能是任何数量的某类用户终端或用户．当给定类中的任一个用户或用户终端向代理服务器发送 URL 时，代理服务器内的处理器访问相关数据库并判定该特定 URL 是否表示对所标记类中的用户/用户终端是允许的请求．图 2 示出了本发明的另一个实施例，它与图 1 所示系统相似，但实现了用户/用户终端类的识别．如图所示，图 2 的系统包括公共网 200，网络资源 201 － 205，用户终端 207 － 210，LAN 211，处理器 212，代理服务器 213 和防火墙 214．图 2 系统的操作基本上与图 1 的相似，除了把两个用户终端 207 和 208 分在一个类中外．这种分组体现在相关数据库 215 的结构中．在相关数据库 215 中，标识码 $ID_{207/208}$ 表示同时与用户终端 207 和用户终端 208 相关．当处理器 212 接收来自用户终端 207 或 208 的 URL 时，将访问与 URL 相关的同一列表．即这两个终端将同

样地被允许或拒绝对同一组 URL（ URL_{101}、 URL_{102} 和 URL_{103} ）的访问.

采用本发明的系统中所用的相关数据库也可以是这样构造的，表示允许访问的资源的信息可以按排成与构造成树状结构的资源一致. 相关数据库可以包括一个允许或拒绝某一特定用户或用户组访问的目录和/或子目录标识号的列表. 例如，可以实现这样的系统，其请求的格式是一组由诸如括弧等分组约定和诸如重复和联合操作的特殊符号一起组成的；标准的表达式对本领域技术人员而言是公知的. 标准表达式规则由一个标准表达式和一个包括或不包括一个或多个用户/用户终端的说明组成、判断一个字符串是否匹配一个标准表达式的标准技术可用来判断一个特定 URL 是否匹配一个标准表达式；这种技术是本领域技术人员所熟悉的.

更一般地说， URL http://ourschool. edu/history/*是一个说明在目录 http://ourschool.edu/history 或其子目录树内的所有资源的标准表达式（一个包括与特定学校的历史课有关的信息的资源）. 在这种情况下，标准表达式采用了 UNIX 的外壳(shell)语言所典型使用的匹配符,其中“ * ”表示任何字符串，包括空串. URL http ://ourschool.edu/subject/*answer* 表示在目录 http ://ourschool.edu/subject （或其子目录树）中并在其名字中包括“ answer ”的所有资源. 对“ answer ”资源的访问将仅限于指导老师（即学生不能看 answer（答案））. 为了说明允许学生看“ history ”资源，但不能看“ history answer ”资源，相关数据库将存储下列与学生标识码相关的带表达式规则的字符串:

+ http://ourschool.edu/history/*

-http://ourschool.edu/history/*answer*

其中标记“ + ”表示允许对资源进行访问，“ - ”表示不能访问.

对本发明的又一个修改允许系统可从用户/用户终端接收除 URL 格式以外的请求. 要储存所用特种类型请求格式的信息指示组，仅需修改相关数据库，并与一类特殊用户相联系.

应该理解的是上面描述的系统和方法仅是为了说明本发明的原理，本领域的技术人员能作出各种变化而不脱离本发明的实质和范围。本发明的实质和范围仅由所附的权利要求书所限定。

图 1

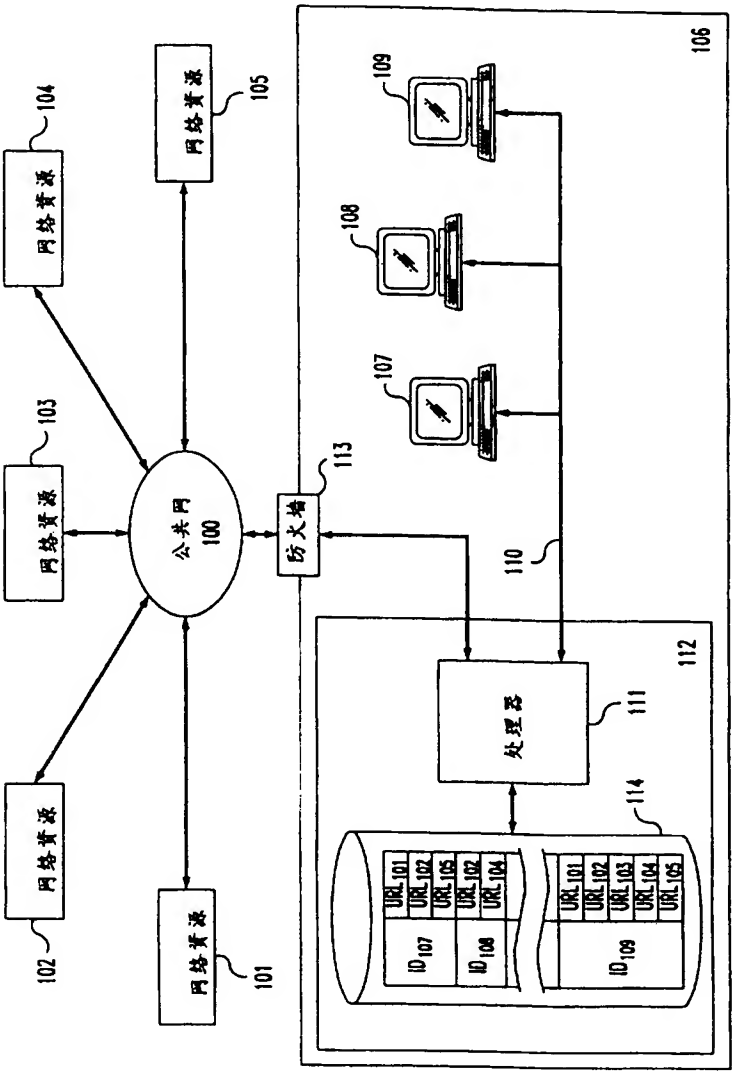
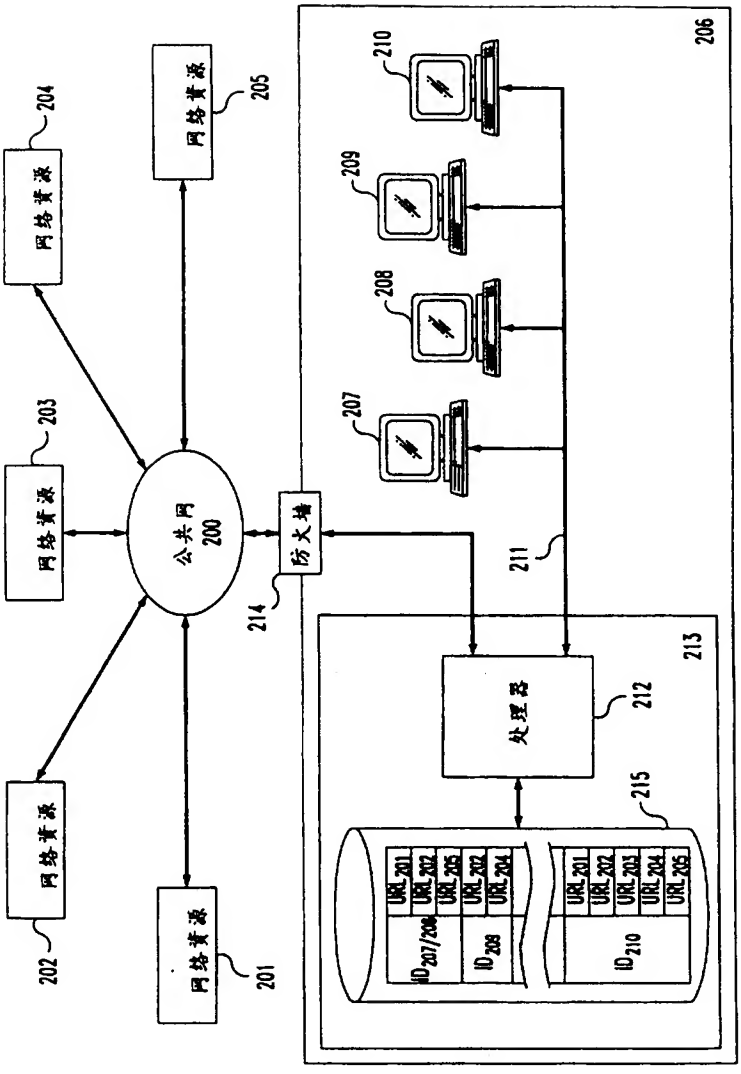

网络资源
101
网络资源
102
网络资源
103
网络资源
104
网络资源
105
公共网
100
防火墙
113
106
109
108
107
110
112
处理器
111
114
URL 101
URL 102
URL 105
URL 102
URL 104
ID 107
ID 108
URL 101
URL 102
URL 103
URL 104
URL 105
ID 109

图 2


网络资源
网络资源
网络资源
网络资源
网络资源
公共网
200
防火墙
处理器
206
210
209
208
207
211
213
212
215
214
URL 201
URL 202
URL 205
URL 202
URL 204
ID 207/208
ID 209
URL 201
URL 202
URL 203
URL 204
URL 205
ID 210